Chapter 6

Differentiation of Measures and Functions

This chapter is concerned with the differentiation theory of Radon measures. In the first two sections we introduce the Radon measures and discuss two covering theorems, the main technical tools in differentiation theory. The Vitali's covering theorem applies to the Lebesgue measure, while the more sophisticated Besicovitch's covering theorem applies to Radon measures. In Section 3 we prove a version of Radon-Nikodym theorem for Radon measures. It differs from the version in Chapter 5 for now there is a good description of the Radon-Nikodym derivative. As application we deduce Lebsegue-Besicovitch differentiation theorem in Section 4. Next we study the differentiability properties of functions in $\mathbb{R}$. First, we show how to relate increasing functions and Radon measures on the real line in Section 5. In Section 6 the concept of functions of bounded variations and absolutely continuous functions are recalled. The former is identified to be those integrable functions whose weak derivatives are signed Radon measures and the latter are those whose weak derivatives are integrable functions.

6.1 Radon Measures

Recall that Riesz representation theorem asserts that for a positive linear functional Λ on the space $C_c(X)$ where X is a locally compact Hausdorff space, there exists an Borel outer measure λ satisfying

$$\Lambda f = \int f d\lambda,$$

for all $f \in C_c(X)$. We used to call λ the Riesz measure associated to Λ. Furthermore, λ enjoys the following regularity properties:

(a) $\lambda(E) = \inf \left\{ \lambda(G) : \quad E \subset G,\ G \text{ open} \right\}$ for every $E \subset \mathbb{R}^n$.

(b) $\lambda(E) = \sup\left\{\lambda(K) : \quad K \subset E,\ K \text{ compact}\right\}$ for every λ-measurable $E \subset \mathbb{R}^n$.

Now, an outer measure μ in $\mathbb{R}^n$ is called a **Radon measure** if

(a) It is **Borel regular**, that is, for every $A \subset \mathbb{R}^n$, there is a Borel set B, $A \subset B$, such that $\mu(A) = \mu(B)$, and

(b) it is finite on compact sets.

Given a Borel measure μ, we define

$$\Lambda f = \int f d\mu, \qquad f \in C_c(X).$$

When μ is finite on compact sets, Λ is a well-defined positive functional on $C_c(X)$. By Riesz representation theorem there is an outer Borel measure λ satisfying

$$\int f d\lambda = \int f d\mu, \qquad \forall f \in C_c(\mathbb{R}^n).$$

For every compact set K sitting inside an open set G, there exists a continuous function f compactly supported in G, equals to 1 in K and bounded between 0 and 1. Plugging such functions in the relation above and passing limit by Lebsegue's dominated convergence theorem, we see that λ and μ coincide on open sets. Consequently, they are the same on all Borel sets. Now, let A be an arbitrary set in $\mathbb{R}^n$. By Borel regularity, there is a Borel set B containing A such that $\mu(B) = \mu(A)$. Therefore,

$$\begin{aligned} \mu(A) &= \mu(B) \\ &= \lambda(B) \\ &\geq \lambda(A) . \end{aligned}$$

On the other hand, by the regularity property of the Riesz measure, for every A and $\varepsilon > 0$, there is an open set G containing A such that $\lambda(A) + \varepsilon \geq \lambda(G)$. Therefore,

$$\begin{aligned} \lambda(A) + \varepsilon &\geq \lambda(G) \\ &= \mu(G) \\ &\geq \mu(A) , \end{aligned}$$

which implies $\lambda(A) \geq \mu(A)$. Summing up, we have proved that every Radon measure is a Riesz measure in $\mathbb{R}^n$.

6.2 Covering Theorems

In the following a non-degenerate closed ball means $\overline{B}_r(x) = \{y \in R^n : \ |y - x| \leq r$ for some $x \in R^n$ and $r > 0$. We will use $\overline{B}$ to denote a non-degenerate closed ball.

Theorem 6.1 (Vitali's Covering Theorem). *Let $\mathcal{F}$ be a collection of non-degenerate closed balls whose diameters are bounded uniformly. There exists a countable subcollection $\mathcal{F}'$ of $\mathcal{F}$ consisting of disjoint balls such that*

$$\bigcup_{\mathcal{F}} \overline{B} \subset \bigcup_{B \in \mathcal{F}'} \hat{B},$$

where $\hat{B}$ is the "5 times" of $\overline{B}$.

The "5 times" of $\overline{B}_r(x)$ is $\overline{B}_{5r}(x)$. Note that the uniform boundedness condition on the diameters cannot be removed. Take $\mathbb{R}$ and $\mathcal{F} = \{[-\alpha, \alpha] : \alpha > \prime\}$, no such $\mathcal{F}'$ can be found.

Proof. Assume that all balls in $\mathcal{F}$ are confined to a bounded set. For each $n \geq 1$, we decompose $\mathcal{F}$ into $\mathcal{F}_n$ according to the diameter of the balls: $\overline{B}$ belongs to $\mathcal{F}_n$ if and only if its diameter falls in $(2^{-n}\rho, 2^{-n+1}\rho]$, where ρ is a bound on the diameter of the balls in $\mathcal{F}$. Since $\mathcal{F}_1$ is bounded, there is an upper bound on the number of disjoint balls in $\mathcal{F}_1$ that can be chosen. Beginning from a single ball, we can add more and more disjoint balls to obtain a largest collection $\mathcal{F}_1'$ which contains finitely many disjoint balls from $\mathcal{F}_1$. Next we consider $\mathcal{F}_2$ after excluding those intersect some ball from $\mathcal{F}_1'$. Repeating the same reasoning we obtain a maximal finite collection $\mathcal{F}_2'$ of disjoint balls. In this way $\mathcal{F}_n'$ are selected for all $n \geq 1$. Then

$$\mathcal{F}' \equiv \bigcup_n \mathcal{F}_n' \ ,$$

is a subcollection of $\mathcal{F}$ consisting of countable, pairwise disjoint balls. Moreover, for any $\overline{B} \in \mathcal{F}_n$, either $\overline{B}$ intersects some ball in $\mathcal{F}_n'$ or it intersects some ball from $\mathcal{F}_m'$, $m \leq n-1$. Letting r and r_1 be the respective radii of $\overline{B}$ and the ball $\overline{B}(z)$ it intersects, $\overline{B}$ is contained in $\overline{B}_{r_1+2r}(z)$. We have

$$\begin{aligned} r_1 + 2r &\leq r_1 + \frac{\rho}{2^{n-1}} \\ &\leq r_1 + 2 \times \frac{\rho}{2^n} \\ &\leq r_1 + 2 \times \frac{\rho}{2^m} \\ &\leq 5r_1 \ . \end{aligned}$$

We conclude that

$$\bigcup_{\mathcal{F}} \overline{B} \subset \bigcup_{\mathcal{F}'} \hat{B}.$$

When $\mathcal{F}$ is not confined to a bounded set, we decompose it as a disjoint union

$$\mathcal{F}_n = \bigcup_k \mathcal{F}_{n,k}, \quad \mathcal{F}_{n,k} = \mathcal{F}_n \bigcap B_k(0) \ .$$

To pick $\mathcal{F}_1'$ we first select a maximal finite collection from $\mathcal{F}_{1,1}$. Next consider those balls in $\mathcal{F}_{1,2}$ disjoint from this maximal finite collection and pick a maximal finite collection. By repeating this procedure we end up with a countable, maximal collection of balls from $\mathcal{F}$ called $\mathcal{F}_1'$. Next we select $\mathcal{F}_n'$ by a similar procedure making sure that the balls now are disjoint from $\mathcal{F}_m', 1 \le m \le n-1$. Then $\mathcal{F}'$, the union of all these $\mathcal{F}_n'$, satisfies our requirement. □

Corollary 6.2. *Let A be a set in $\mathbb{R}^n$ and $\mathcal{F}$ a family of nondegenerate, closed balls with uniformly bounded diameters that covers A in the following sense: For each $x \in A$, there is a sequence of balls centered at x with radii going to 0 from $\mathcal{F}$. For every open set G containing A, there is a countable subfamily $\mathcal{F}'$ consisting of disjoint balls from $\mathcal{F}$ such that*

$$\bigcup_{\mathcal{F}'} \overline{B} \subset G ,$$

and

$$\mathcal{L}^n \left(A \setminus \bigcup_{\mathcal{F}'} \overline{B} \right) = 0 .$$

Proof. Assume first that A is bounded. By the outer regularity of the Lebesgue measure, we can fix an open set $G_1 \subset G$ containing A such that $\mathcal{L}^n(G_1) \le (1+\varepsilon_0)\mathcal{L}^n(A)$ where ε_0 satisfies

$$\theta \equiv 1 + 2\varepsilon_0 - \frac{1}{5^n} \in (0,1) .$$

Let

$$\mathcal{F}_1 = \left\{ \overline{B}(x) : \ x \in A, \ \overline{B}(x) \in \mathcal{F}, \ \overline{B}(x) \subset G_1 \right\} .$$

Applying Vitali's covering theorem, there are pairwise disjoint balls $B_k, k \ge 1$, from $\mathcal{F}_1$ that satisfy

$$A \subset \bigcup_{\mathcal{F}_1} \overline{B} \subset \bigcup_k \hat{B}_k .$$

We have

$$
\begin{aligned}
\mathcal{L}^n\left(A\setminus\bigcup_k \overline{B}_k\right) &\le \mathcal{L}^n\left(G_1\setminus\bigcup_k \overline{B}_k\right)\\
&= \mathcal{L}^n(G_1)-\mathcal{L}^n\left(\bigcup_k \overline{B}_k\right)\\
&\le \mathcal{L}^n(G_1)-\sum_k \mathcal{L}^n(\overline{B}_k)\\
&= \mathcal{L}^n(G_1)-\frac{1}{5^n}\sum_k \mathcal{L}^n(\hat{B}_k)\\
&\le (1+\varepsilon_0)\mathcal{L}^n(A)-\frac{1}{5^n}\mathcal{L}^n\left(\bigcup_k \hat{B}_k\right)\\
&\le (1+\varepsilon_0-\frac{1}{5^n})\mathcal{L}^n(A)\\
&< \theta\mathcal{L}^n(A)\ .
\end{aligned}
$$

We can fix a large m so that

$$\mathcal{L}^n\left(A\setminus\bigcup_{k=1}^m \overline{B}_k\right)\le \theta\mathcal{L}^n(A)\ .$$

Next choose an open set $G_2\subset G_1$ containing

$$A_2\equiv A\setminus\bigcup_{k=1}^m \overline{B}_k$$

such that

$$\mathcal{L}^n(G_2)\le(1+\varepsilon_0)\mathcal{L}^n(A_2)\ ,$$

with the same ε_0 specified before. Let

$$\mathcal{F}_2=\left\{\overline{B}(x):\ x\in A_2,\ \overline{B}(x)\in\mathcal{F},\ \overline{B}(x)\cap\overline{B}_k=\phi,\ k=1,\cdots,m,\ \overline{B}(x)\subset G_2\right\}\ .$$

Since $\cup_{k=1}^m\overline{B}_k$ is closed, every point in A_2 belongs to some ball in $\mathcal{F}_2$. We have

$$A_2\subset\bigcup_{\mathcal{F}_2}\overline{B}\ .$$

Repeat the argument to (A_2,G_2) instead of (A,G_1) we obtain finitely many pairwise disjoint balls from $\mathcal{F},\overline{B}_{m+1},\cdots,\overline{B}_l$ such that

$$\mathcal{L}^n\left(A_2\setminus\bigcup_{k=m+1}^{l}\overline{B}_k\right)\le\theta\mathcal{L}^n(A_2)\ .$$

In other words,

$$\mathcal{L}^n\left(A \setminus \bigcup_{k=1}^{l} \overline{B}_k\right) \le \theta^2 \mathcal{L}^n(A) \ .$$

Repeating this argument, we come up with a countable, pairwise disjoint balls from $\mathcal{F}$ such that

$$\begin{aligned} \mathcal{L}^n\left(A \setminus \bigcup_{k} \overline{B}_k\right) &\le \liminf_{p\to\infty} \mathcal{L}^n\left(A \setminus \bigcup_{k=1}^{p} \overline{B}_k\right) \\ &\le \liminf_{p\to\infty} \theta^p \mathcal{L}^n(A) \\ &= 0 \ . \end{aligned}$$

In general, for an unbounded A, let

$$R_n = \{n < |x| < n+1\} , \quad A_n = A \bigcap R_n, \ n \ge 1,$$

and $A_0 = A \bigcap B_1(0)$ and apply the previous result to each $A_n, \ n \ge 0$. Note that

$$\bigcup_n \{x : |x| = n\} \cap A$$

is a null set.

□

Vitali's covering theorem is effective for the Lebesgue measure. Unfortunately, for an arbitrary Radon measure, there may be no relation between the measure of the "5 times" of a ball to the measure of the ball. In this case, we need a more powerful covering result.

Theorem 6.3 (**Besicovitch's Covering Theorem**). *Let $\mathcal{F}$ be a collection of nondegenerate closed balls whose diameters are uniformly bounded and let A be the set consisting of the centers of these balls. There exists finitely many subcollections $\mathcal{F}_1, \dots, \mathcal{F}_N$ of $\mathcal{F}$ in which each $\mathcal{F}_j$ is composed of countably many disjoint balls such that*

$$A \subset \bigcup_{j=1}^{N} \bigcup_{\mathcal{F}_j} \overline{B}.$$

The number N depends only on the dimension n.

Corollary 6.4. *Let μ be a Borel measure and A a set in $\mathbb{R}^n$ with $\mu(A) < \infty$. Let $\mathcal{F}$ a family of nondegenerate, closed balls with uniformly bounded diameters that covers A in the following sense: For each $x \in A$, there is a sequence of balls centered at x with radii going to 0 from $\mathcal{F}$. For every open set G containing A, there is a countable subfamily $\mathcal{F}'$ from $\mathcal{F}$ such that*

$$\bigcup_k \overline{B} \subset G \ ,$$

and

$$\mu\left(A \setminus \bigcup_{\mathcal{F}'} \overline{B}\right) = 0 \ .$$

The proof of Corollary 6.4 is very much like that of Corollary 6.2 where the only difference is to replace Vitali's theorem by Besicovitch's theorem. We refer to [EG] for the details.

6.3 Radon-Nikodym Derivatives

Let μ and ν be two Radon measures on $\mathbb{R}^n$. We define the lower and upper derivatives of ν with respect to μ by

$$\underline{D}_\mu \nu(x) = \begin{cases} \displaystyle\varliminf_{r\to 0} \frac{\nu(\overline{B}_r(x))}{\mu(\overline{B}_r(x))}, & \mu(\overline{B}_r(x)) > 0, \ \forall r > 0 \\ \infty, & \mu(\overline{B}_r(x)) = 0, \ \text{some } r > 0. \end{cases}$$

and

$$\overline{D}_\mu \nu(x) = \begin{cases} \displaystyle\varlimsup_{r\to 0} \frac{\nu(\overline{B}_r(x))}{\mu(\overline{B}_r(x))}, & \mu(\overline{B}_r(x)) > 0, \ \forall r > 0 \\ \infty, & \mu(\overline{B}_r(x)) = 0, \ \text{some } r > 0. \end{cases}$$

Here $\underline{D}_\mu \nu$ and $\overline{D}_\mu \nu$ are functions from $\mathbb{R}^n$ to $[0, \infty]$.

Lemma 6.5. *Let μ and ν be Radon measures on $\mathbb{R}^n$. For every set A and $\alpha \in (0, \infty)$,*

(a) $A \subset \{x : \underline{D}_\mu \nu(x) \le \alpha\}$ implies that $\nu(A) \le \alpha\mu(A)$.

(b) $A \subset \{x : \overline{D}_\mu \nu(x) \ge \alpha\}$ implies that $\nu(A) \ge \alpha\mu(A)$.

Proof. We only prove (a). Assume first that A is bounded. For $\varepsilon > 0$, we pick an open G containing A. For each $x \in A$ there is a sequence of closed balls $\overline{B}$ in G centering at x with radius less than 1 and going down to 0 such that $\nu(\overline{B}) \le (\alpha+\varepsilon)\mu(\overline{B})$. The collection of all these balls from a covering of A. Applying Corollary 6.4 to ν, there is a countable subcollection consisting of disjoint closed balls, $\overline{B}_j, j \ge 1$, such that $\nu(A \setminus \bigcup_j \overline{B}_j) = 0$. It follows that

$$\begin{aligned}
\nu(A) &\leq \nu\left(A \setminus \bigcup_j \overline{B}_j\right) + \nu\left(\bigcup_j \overline{B}_j\right) \\
&= \nu\left(\bigcup_j \overline{B}_j\right) \\
&= \sum_j \nu(\overline{B}_j) \\
&\leq (\alpha+\varepsilon)\sum_j \mu(\overline{B}_j) \\
&= (\alpha+\varepsilon)\mu\left(\bigcup_j \overline{B}_j\right) \\
&\leq (\alpha+\varepsilon)\mu(G).
\end{aligned}$$

Taking infimum over all G, by the regularity property of μ we get $\nu(A) \leq (\alpha+\varepsilon)\mu(A)$ which implies the desired result after letting ε go to 0. When A is unbounded, apply the previous proof to $A \cap B_n(0)$ and then let n go to ∞. □

The technique of proof of this lemma is standard and you should understand it well.

Remark 6.1. When μ is the Lebsegue measure and $\nu << \mathcal{L}^n$, Lemma 6.5 can be proved by Corollary 6.2 instead of Corollary 6.4. In other words, Besicovitch's covering theorem could be replaced by Vitali's covering theorem in this special case. See exercise.

Theorem 6.6. *Let μ and ν be Radon measures on $\mathbb{R}^n$. Then*

(i) The Borel set $\{x : \overline{D}_\mu\nu(x) = \infty\}$ is μ-null.

(ii) The Borel set $\{x : \underline{D}_\mu\nu(x) < \overline{D}_\mu\nu(x) < \infty\}$ is μ-null.

Note that $\overline{D}_\mu\nu(x) \geq \underline{D}_\mu\nu(x)$, $\forall x$, and $\underline{D}_\mu\nu(x) \geq 0$ from definition.

Proof. For a fixed k, $\{x : |x| \leq k,\ \overline{D}_\mu\nu(x) = \infty\}$ is contained in $\{x : |x| \leq k, \overline{D}_\mu\nu(x) \geq \alpha\}$ for all $\alpha > 0$. By Lemma 6.5(ii),

$$\begin{aligned}
\mu\left(\{x : |x| \leq k,\ \overline{D}_\mu\nu(x) = \infty\}\right) &\leq \frac{1}{\alpha}\nu\left(\{x : |x| \leq k, \overline{D}_\mu\nu(x) \geq \alpha\}\right) \\
&\leq \frac{1}{\alpha}\nu(\overline{B}_k)\ ,
\end{aligned}$$

which tends to 0 as $\alpha \to \infty$.

Next, for $0 < \alpha < \beta$, let $A(\alpha,\beta) \equiv \left\{x : |x| \le k,\ \underline{D}_\mu\nu(x) < \alpha < \beta < \overline{D}_\mu\nu(x) < \infty\right\}$. By Lemma 6.5,

$$\nu(A(\alpha,\beta)) \le \alpha\mu(A(\alpha,\beta)) , \quad \beta\mu(A(\alpha,\beta)) \le \nu(A(\alpha,\beta)) ,$$

which forces $\mu(A(\alpha,\beta)) = 0$. Observing that

$$\left\{x : |x| \le k,\ \underline{D}_\mu\nu(x) < \overline{D}_\mu\nu(x) < \infty\right\} = \bigcup_{\alpha,\beta} A(\alpha,\beta) ,$$

which $\alpha < \beta$ run over all rational numbers, we conclude that

$$\mu\left(\left\{x : |x| \le k,\ \underline{D}_\mu\nu(x) < \overline{D}_\mu\nu(x) < \infty\right\}\right) = 0 .$$

□

This theorem shows that $\overline{D}_\mu\nu = \underline{D}_\mu\nu < \infty$ except on a μ-measure zero set N where $N = I \bigcup R$,

$$I = \left\{x : \overline{D}_\mu\nu(x) = \infty\right\} , \quad R = \left\{x : \underline{D}_\mu\nu(x) < \overline{D}_\mu\nu(x) < \infty\right\} .$$

To make the derivative well-defined everywhere, we set

$$\begin{aligned} D_\mu\nu(x) &= \overline{D}_\mu\nu(x), \quad && x \in \mathbb{R}^n \setminus N \\ &= \infty, && x \in N. \end{aligned}$$

Proposition 6.7. *Let μ and ν be Radon measures on $\mathbb{R}^n$. Then $D_\mu\nu$ is μ-measurable.*

Proof. We claim that $x \mapsto \mu(\overline{B}_r(x))$ is upper semicontinuous. Let $x_n \to x$. For a fixed $\varepsilon > 0$, $\overline{B}_r(x_n)$ is contained in $\overline{B}_{r+\varepsilon}(x)$ for all large n. It follows that

$$\mu(\overline{B}_r(x_n)) \subset \mu(\overline{B}_{r+\varepsilon}(x) ,$$

and

$$\overline{\lim_{n\to\infty}}\, \mu(\overline{B}_r(x_n)) \le \mu(\overline{B}_{r+\varepsilon}(x)) .$$

Now, letting $\varepsilon \to 0$, by the monotone convergence theorem

$$\overline{\lim_{n\to\infty}}\, \mu(\overline{B}_r(x_n)) \le \mu(\overline{B}_r(x)) ,$$

which means that $x \mapsto \mu(\overline{B}_r(x))$ is upper semicontinuous. Similarly one can show that $x \mapsto \nu(\overline{B}_r(x))$ is upper semicontinuous. For $x \in \mathbb{R}^n \setminus N$, noting that $\mu(\overline{B}_r(x)) > 0$ for all r, we can fix a sequence $r_k \to 0$ to get

$$\begin{aligned} D_\mu\nu(x) &= \lim_{r\to 0} \frac{\nu(\overline{B}_r(x))}{\mu(\overline{B}_r(x))} \\ &= \lim_{r_k\to 0} \frac{\nu(\overline{B}_r(x))}{\mu(\overline{B}_r(x))} < \infty . \end{aligned}$$

Since N is a μ-measurable null set where $D_\mu\nu = \infty$, we conclude that $D_\mu\nu$ is μ-measurable.

□

We have the following version of Radon-Nikodym theorem. Comparing with the Radon-Nikodym theorem we proved in Chapter 5, the main difference is that we have a more concrete description of the Radon-Nikodym derivative here.

Theorem 6.8 (Radon-Nikodym Theorem). *Let μ and ν be Radon measures on $\mathbb{R}^n$ satisfying $\nu \ll \mu$. For every μ-measurable A,*

$$\nu(A) = \int_A D_\mu \nu \, d\mu.$$

Proof. Assume that A is bounded. We first claim that A is μ-measurable implies that A is ν-measurable too. For, as μ is a Radon measure, we can find a Borel set (in fact, a G_δ-set) A_1, $A \subset A_1$, such that $\mu(A_1 \setminus A) = 0$. As $\mu \gg \nu$, $\nu(A_1 \setminus A) = 0$. Using the fact that every ν-null set is ν-measurable, $A = A_1 \setminus (A_1 \setminus A)$ is ν-measurable.

Consider the "exceptional sets"

$$\begin{aligned} I &= \left\{x : \overline{D}_\mu \nu = \infty\right\}, \\ Z &= \left\{x : D_\mu \nu = 0\right\}, \text{ and} \\ R &= \left\{x : \underline{D}_\mu \nu < \overline{D}_\mu \nu < \infty\right\}. \end{aligned}$$

We know that $\mu(I) = \mu(R) = 0$ from Lemma 6.5 and Theorem 6.6 respectively. Although $\mu(Z)$ may not vanish, $\nu(Z) = 0$ by Lemma 6.5 too.

For a μ-measurable A, let

$$A_m = \left\{x \in A : t^m < D_\mu \nu(x) \le t^{m+1}\right\}, \quad m \in \mathbb{Z},$$

where $t > 1$ is fixed. We have

$$A \setminus \bigcup_{-\infty}^{\infty} A_m \subset I \cup Z \cup R.$$

Therefore,

$$\nu\left(A \setminus \bigcup_{-\infty}^{\infty} A_m\right) = 0, \quad \text{or} \quad \nu(A) = \sum_{-\infty}^{\infty} \nu(A_m),$$

since $\nu(I) = \nu(R) = 0$ by $\mu(I) = \mu(R) = 0$ and $\nu \ll \mu$. Note that A and A_m are

μ-measurable implies that they are also ν-measurable. We have

$$\begin{aligned}
\nu(A) &= \nu\left(\bigcup_{-\infty}^{\infty} A_m\right) \\
&= \sum_{-\infty}^{\infty} \nu(A_m) \\
&\le \sum_{-\infty}^{\infty} t^{m+1}\mu(A_m) \quad (\text{ use } A_m \subset \left\{x : D_\mu\nu(x) \le t^{m+1}\right\} \text{ and Lemma 6.5.)} \\
&= t\sum_{-\infty}^{\infty} t^m \mu(A_m) \\
&\le t\sum_{-\infty}^{\infty} \int_{A_m} D_\mu\nu\, d\mu \\
&\le t\int_{\bigcup A_m} D_\mu\nu\, d\mu \\
&\le t\int_A D_\mu\nu\, d\mu.
\end{aligned}$$

Letting $t \downarrow 1$,

$$\nu(A) \le \int_A D_\mu\nu\, d\mu.$$

On the other hand,

$$\begin{aligned}
\nu(A) &= \sum_{-\infty}^{\infty} \nu(A_m) \\
&\ge \frac{1}{t}\sum_{-\infty}^{\infty} t^{m+1}\nu(A_m) \quad (\text{ use } \ A_m \subset \{x : t^m < D_\mu\nu(x)\} \text{ and Lemma 6.5.)} \\
&\ge \frac{1}{t}\sum_{-\infty}^{\infty} \int_{A_m} D_\mu\nu\, d\mu \\
&= \frac{1}{t}\int_{\bigcup A_m} D_\mu\nu\, d\mu \\
&= \frac{1}{t}\int_A D_\mu\nu\, d\mu \quad \Bigg(\text{use } \ A \setminus \bigcup_m A_m \subset I \cup Z \cup R \text{ and} \\
&\qquad\qquad \int_I D_\mu\nu\, d\mu = \int_R D_\mu\nu\, d\mu = \int_Z D_\mu\nu\, d\mu = 0\Bigg)
\end{aligned}$$

Letting $t \uparrow 1$,

$$\nu(A) \ge \int_A D_\mu\nu\, d\mu.$$

Finally, when A is unbounded, apply the above result to $A \cap B_n(0)$ and then let $n \to \infty$. □

Remark 6.2. From Remark 6.1 we see that when μ is the Lebesgue measure, the proof of the Radon-Nikodym theorem can be based on Vitali's covering theorem and is independent of Besicovitch's covering theorem.

In Theorem 6.9, the measure ν is absolutely continuous with respect to μ. More generally, for two unrelated Radon measures we have

Theorem 6.9. *Let μ and ν be two Radon measures on $\mathbb{R}^n$. There exists a Borel set $A^* \subset \mathbb{R}^n$, $\mu(\mathbb{R}^n \setminus A^*) = 0$, such that by setting*

$$\nu_{ac} = \nu\lfloor_{A^*} \qquad \text{and} \qquad \nu_s = \nu\lfloor_{\mathbb{R}^n\setminus A^*} ,$$

we have $\nu_{ac} \ll \mu$ and $\nu_s \perp \mu$. Moreover, $D_\mu \nu_s = 0$ μ-a.e. so that $D_\mu \nu = D_\mu \nu_{ac}$ μ-a.e.

The measures ν_{ac} and ν_s are Radon. Recall that this decomposition is called the Lebesgue decomposition of ν with respect to μ. We learnt it in Chapter 5, but here we give a second proof which contains the additional information $D_\mu \nu_s = 0$.

Proof. Let

$$\mathcal{E} = \{A \subset \mathbb{R}^n : \mu(\mathbb{R}^n \setminus A) = 0,\ A \text{ is a Borel set}\} .$$

We claim there exists an $A^* \in \mathcal{E}$ such that

$$\nu(A^*) \le \nu(A), \quad \forall A \in \mathcal{E}.$$

For, let $A_k \in \mathcal{E}$ satisfy

$$\nu(A_k) \le \inf_{\mathcal{E}} \nu(A) + \frac{1}{k}$$

and $A^* = \bigcap A_k$. Clearly, $A^* \in \mathcal{E}$ and $\nu(A^*) \le \nu(A)$.

Define ν_{ac} and ν_s as in the theorem. Clearly, $\nu_s \perp \mu$. To show $\nu_{ac} \ll \mu$, note that $E = (E \cap A^*) \cup (E \setminus A^*)$. As $\nu_{ac}(E \setminus A^*) = 0$ by definition, it suffices to assume that $E \subset A^*$ and check $\mu(E) = 0$ implies that $\nu_{ac}(E) = 0$. Suppose on the contrary there is some $E_1 \subset A^*$ such that $\mu(E_1) = 0$ but $\nu_{ac}(E_1) > 0$. Find a Borel set E_2, $E_1 \subset E_2$, such that $\mu(E_2) = 0$ and $\nu_{ac}(E_2) = \nu_{ac}(E_1) > 0$. The set $A^* \setminus E_2 \in \mathcal{E}$ but

$$\begin{aligned} \nu_{ac}(A^* \setminus E_2) &= \nu_{ac}(A^*) - \nu_{ac}(E_2) \\ &< \nu_{ac}(A^*) \\ &= \inf_{\mathcal{E}} \nu(A^*), \end{aligned}$$

contradiction holds. Hence $\nu_{ac} \ll \mu$.

To show that $D_\mu \nu_s = 0$ μ-a.e., let $E = \{x : \ D_\mu \nu_s \geq \alpha\}$ where $\alpha > 0$ is fixed. By Lemma 6.5 we have $\nu_s(E \cap A^*) \geq \alpha \mu(E \cap A^*)$. As $\nu_s(A^*) = 0$, we have $\mu(E \cap A^*) = 0$, but then $\mu(E) = \mu(E \cap A^*) + \mu(E \cap (\mathbb{R}^n \setminus A^*) = 0$, that is, $D_\mu \nu_s = 0$ μ-a.e.. It follows that $D_\mu \nu = D_\mu \nu_{ac} + D_\mu \nu_s = D_\mu \nu_{ac} = 0$ μ-a.e.. □

6.4 Lebesgue Points

Theorem 6.10 (**Lebesgue's Differentiation Theorem**). *Let μ be a Radon measure on $\mathbb{R}^n$. For $f \in L^1_{loc}(\mu)$,*

$$\lim_{r \to 0} \frac{1}{\mu(\overline{B}_r(x))} \int_{\overline{B}_r(x)} f \, d\mu = f(x),$$

for μ-a.e. x.

This result is due to Lebesgue for $\mu = \mathcal{L}^1$ and the general case is due to Besicovitch. It is also called **Lebesgue-Besicovitch differentiation theorem.**

Proof. Clearly it is sufficient to prove the theorem by assuming $f \geq 0$ and $f \in L^1(\mathbb{R}^n)$. Define a positive linear functional Λ be

$$\Lambda \varphi = \int \varphi f d\mu, \quad \varphi \in C_c(\mathbb{R}^n) \ .$$

By Riesz representation theorem, there is a Radon measure ν satisfying

$$\int \varphi d\nu = \int \varphi f d\mu \ .$$

By a routine argument,

$$\nu(E) = \int_E f \, d\mu, \quad E \ \mu\text{-measurable} \ .$$

By Theorem 6.6,

$$\lim_{r \to 0} \frac{1}{\mu(\overline{B}_r(x))} \int_{\overline{B}_r(x)} f(y) \, d\mu(y) = D_\mu \nu(x),$$

for μ-a.e. x. As $\nu << \mu$, by Theorem 6.8,

$$\nu(E) = \int_E f d\mu = \int_E D_\mu \nu \, d\mu, \quad E \ \mu\text{-measurable} \ ,$$

hence $f = D_\mu \nu$ μ-a.e.. Combining these we get the theorem. □

The corollary gives a stronger result.

Corollary 6.11. *Let $f \in L^1_{loc}(\mu)$ where μ is a Radon measure. There exists a set L of full measure such that for every $x \in L$, $f(x)$ is finite and*

$$\lim_{r\to 0} \frac{1}{\mu(\overline{B}_r(x))} \int_{\overline{B}_r(x)} |f - f(x)| \, d\mu = 0.$$

Proof. Order all rational numbers in a sequence $\{q_j\}$ and apply Theorem 6.10 to each $g_j = |f - q_j|$ to get

$$\frac{1}{\mu(\overline{B}_r(x))} \int_{\overline{B}_r(x)} |f(y) - q_j| \, d\mu(y) \to |f(x) - q_j| \,, \quad \forall x \notin N_j, \tag{6.1}$$

as $r \to 0$ where $\mu(N_j) = 0$. Letting $N = \bigcup_{j=1}^{\infty} N_j$, $\mu(N) = 0$ and (6.1) holds for all $x \notin N$.

Given $\varepsilon > 0$ and $x \notin N$ such that $f(x)$ is finite, pick q_j such that $|f(x) - q_j| < \varepsilon$. We have

$$\begin{aligned}\frac{1}{\mu(\overline{B}_r(x))} \int_{\overline{B}_r(x)} |f(y) - f(x)| \, d\mu(y) &\le \frac{1}{\mu(\overline{B}_r(x))} \int_{\overline{B}_r(x)} |f(y) - q_j| \, d\mu(y) \\ &\quad + \frac{1}{\mu(\overline{B}_r(x))} \int_{\overline{B}_r(x)} |q_j - f(x)| \, d\mu(y) \\ &= \frac{1}{\mu(\overline{B}_r(x))} \int_{\overline{B}_r(x)} |f(y) - q_j| \, d\mu(y) + |q_j - f(x)| \,.\end{aligned}$$

By (6.1), for $\varepsilon > 0$, we can find some r_0 such that for all $r \in (0, r_0)$,

$$\frac{1}{\mu(\overline{B}_r(x))} \int_{\overline{B}_r(x)} |f(y) - q_j| \, d\mu(y) < \varepsilon.$$

Therefore, we have

$$\frac{1}{\mu(\overline{B}_r(x))} \int_{\overline{B}_r(x)} |f - f(x)| \, d\mu(y) \le 2\varepsilon.$$

□

A point x at which this corollary holds is called a **Lebesgue point** for the function f and the set of all Lebesgue points forms the **Lebesgue set** of f. It is implicitly assumed that the function is finite at a Lebesgue point. Moreover, the Lebesgue point depends on the pointwise definition of f and therefore is not a concept attached to f as an equivalence class. Every point of continuity of f is a Lebesgue point, but the converse may be not true. It is not hard to construct Lebesgue points at which the function is discontinuous.

So far we have been considering taking the average of a function over balls. Now we consider taking average over other sets. Let $\{E_j\}$ be a sequence of μ-measurable sets. We call it **shrinks regularly** to x if there is $\overline{B}_{r_j}(x)$, $r_j \to 0$, such that

$$\alpha\mu(E_j) \ge \mu(\overline{B}_{r_j}(x)), \quad E_j \subset \overline{B}_{r_j}(x),$$

for some $\alpha > 0$ (depending on x).

Corollary 6.12. *Let $\{E_j\}$ shrink regularly to x, a Lebesgue point of $f \in L^1_{loc}(\mathbb{R}^n)$. Then*

$$\lim_{j\to\infty} \frac{1}{\mu(E_j)} \int_{E_j} f \, d\mu = f(x).$$

Proof. By Corollary 6.11,

$$\begin{aligned} \frac{1}{\mu(E_j)} \int_{E_j} |f - f(x)| \, d\mu &\le \alpha \frac{1}{\mu(\overline{B}_{r_j}(x))} \int_{E_j} |f - f(x)| \, d\mu \\ &\le \alpha \frac{1}{\mu(\overline{B}_{r_j}(x))} \int_{\overline{B}_{r_j}(x)} |f - f(x)| \, d\mu \\ &\to 0 \text{ as } r_j \to 0. \end{aligned}$$

□

Let E be a set in $\mathbb{R}^n$. A point $x \in \mathbb{R}^n$ is called a **point of density** α of E if

$$\lim_{r\to 0} \frac{\mu(E \cap \overline{B}_r(x))}{\mu(\overline{B}_r(x))} \in [0,1] \ ,$$

exists and is equal to α.

Corollary 6.13. *Let E be μ-measurable. Every Lebesgue point of χ_E is a point of density 1. Thus the density is 1 for μ-a.e. $x \in E$.*

Proof. Let $f = \chi_E \in L^1_{loc}(\mu)$ and x a Lebesgue point of f. We have

$$\frac{\mu(E \cap \overline{B}_r(x))}{\mu(\overline{B}_r(x))} = \frac{1}{\mu(\overline{B}_r(x))} \int_{\overline{B}_r(x)} \chi_E \, d\mu \to \chi_E(x) = 1,$$

as $r \to 0$. Every Lebesgue point has density one. □

Note that for x lying outside E, applying the same argument to $\mathbb{R}^n \setminus E$, we conclude that for μ-a.e. points lying out E their density are equal to 0.

6.5 Measures and Distribution Functions

In this section general results obtained in the previous sections are applied to functions on the real line. In particular, we will establish the differentiability of functions of bounded variations and prove the fundamental theorem of calculus for absolutely continuous functions. Since many results have been known from undergraduate real analysis, we will be sketchy here or there.

There is a close relationship between Radon measures and monotone functions in $\mathbb{R}$, and similarly for signed Radon measures and functions of bounded variation.

Let μ be a finite Radon measure on $\mathbb{R}$. Its **distribution function** is given by

$$f(x) = \mu((-\infty, x)).$$

(In some texts, $\mu(\infty, x]$ is used instead, but the difference is insignificant.) It is readily checked that f satisfies the following properties:

(a) f is increasing,

(b) f is left continuous, i.e., $\lim_{y\uparrow x} f(y) = f(x)$,

(c) $f(-\infty) = \lim\limits_{x\to-\infty} f(x) = 0$,

(d) $f(\infty) = \lim\limits_{x\to-\infty} f(x) < \infty$.

Conversely, we have

Proposition 6.14. *Let f be a function on $\mathbb{R}$ satisfying (a)–(d). There exists a unique finite Radon measure μ such that f is the distribution function of μ.*

Here we sketch a proof based on the Riemann-Stieltjes integral.

Let P be a partition of the real line and consider the Riemann-Stieltjes sum for a function φ in $C_c(\mathbb{R})$.

$$\sum_{-\infty}^{\infty} \varphi(z_j)(f(x_{j+1}) - f(x_j)),$$

where $z_j \in [x_j, x_{j+1}]$ and $P : \cdots < x_j < x_{j+1} < x_{j+1} < \cdots$, is the partition. It is not hard to show that as $\|P\| = \sup_j |x_{j+1} - x_j| \to 0$, these sums tend to a definite number called the Riemann-Stieltjes integral of φ, denoted by

$$\int \varphi \, df.$$

From the definition, it is clear the Riemann-Stieltjes integral is a positive linear functional on $C_c(\mathbb{R})$. By Riesz representation theorem, there exists a Radon measure μ such that

$$\int \varphi \, df = \int \varphi \, d\mu.$$

By suitably choosing φ, one can deduce $f(x) = \mu((-\infty, x))$, that is, F is the distribution function for μ.

Let μ_1 be another Radon measure satisfying $f(x) = \mu_1((-\infty, x))$. From $\mu_1((-\infty, x)) = \mu((-\infty, x))$ one deduces $\mu_1((x, y)) = \mu((x, y))$, $\forall x \le y$, and μ_1 coincides with μ.

As an application of the connection between measures and functions we prove the almost everywhere differentiability of monotone functions. First of all, we recall that every monotone function defined on an interval has at most countably many discontinuity points which are jumps. Therefore, by redefining the functions at these points, one obtains a left continuous monotone function which is equal to the original function except possibly at a countable set.

Proposition 6.15. *Every monotone function on $[a, b]$ is differentiable almost everywhere.*

Proof. Let f_0 be increasing on $[a, b]$. Extend it to be constant below a and beyond b so that it is increasing on $\mathbb{R}$. Call the extension f. By the remark above we may assume f is left continuous. Moreover, by subtracting it by a constant, we may assume that it satisfies (a)-(d) above.

Let μ be the finite Radon measure taking f as its distribution function. We have the Lebesgue decomposition $\mu = \mu_{ac} + \mu_s$ with respect to the Lebesgue measure. By Theorem 6.9 $D\mu = D\mu_{ac}$, $D\mu_s = 0$ on a set of full measure E_1. Note that we have dropped the subscript $\mathcal{L}^1$. By Corollary 6.12,

$$\begin{aligned} \lim_{\delta \to 0} \frac{\mu_{ac}([x, x+\delta))}{\delta} &= \lim_{\delta \to 0} \frac{1}{\delta} \int_{[x, x+\delta)} \frac{D\mu_{ac}}{D\mathcal{L}^1} d\mathcal{L}^1 \\ &= \frac{D\mu_{ac}}{D\mathcal{L}^1}(x) \ , \end{aligned}$$

for x in a set of full measure E_2. For $x \in E_1$,

$$\begin{aligned} \overline{\lim_{\delta \to 0}} \frac{\mu_s([x, x+\delta))}{\delta} &\le 2 \overline{\lim_{\delta \to 0}} \frac{\mu_s([x-\delta, x+\delta])}{2\delta} \\ &= 0 \ . \end{aligned}$$

Therefore, for $x \in E_1 \cap E_2$,

$$\lim_{\delta \to 0} \frac{\mu([x, x+\delta))}{\delta} = \frac{D\mu}{D\mathcal{L}^1}(x) \ .$$

Using the relation $\mu([x, x+\delta)) = f(x+\delta) - f(x)$, we conclude that

$$\frac{d^+ f}{dx}(x) = \frac{D\mu}{D\mathcal{L}^1}(x), \quad x \in E_1 \cap E_1 \ .$$

On the other hand, by a similar argument we can show that

$$\frac{d^- f}{dx}(x) = \frac{D\mu}{D\mathcal{L}^1}(x)$$

on a set of full measure. The desired result follows. □

Proposition 6.16. *Let f be the distribution function of the finite Radon measure μ. Then*

$$\int \varphi' f d\mathcal{L}^1 = -\int \varphi d\mu \ , \quad \forall \varphi \in C_c(\mathbb{R}) \cap C^1(\mathbb{R}) \ . \tag{6.2}$$

Conversely, suppose that (6.2) holds for some integrable function f and finite Radon measure μ. Then f admits a representative which is increasing and $f+c$ is the distribution function of μ for some constant c.

When $\mu << \mathcal{L}^1$, $d\mu = f' d\mathcal{L}^1$. Hence (6.2) may be interpreted as the "weak derivative" of f is the measure μ.

Proof. Looking at a sequence of partitions with length δ going to 0, we have

$$\begin{aligned}
\int \varphi d\mu &= \int \varphi df \\
&= \lim_{\delta \to 0} \sum_k \varphi(x_{k+1})(f(x_{k+1}) - f(x_k)) \\
&= -\lim_{\delta \to 0} \sum_k (\varphi(x_{k+1}) - \varphi(x_k)) f(x_k) \\
&= -\lim_{\delta \to 0} \sum_k \varphi'(z_k) f(x_k)(x_{k+1} - x_k) \\
&= -\int \varphi' f d\mathcal{L}^1 \ .
\end{aligned}$$

Conversely, let g be the distribution function of μ. By what we just proved, (6.2) holds when f is replaced by g. It follows that

$$\int \varphi'(f-g) \, d\mu = 0, \quad \forall \varphi \in C_c(\mathbb{R}) \cap C^1(\mathbb{R}) \ .$$

Let x and y be two Lebesgue points of f satisfying $x < y$. For a small $\delta > 0$, let φ be the continuous function which is equal to 1 on $[x, y]$, 0 outside $[x-\delta, y]$ and linear in between. By approximation it is legal for a test function. As a result,

$$\frac{1}{\delta} \int_{x-\delta}^{x} (f-g) d\mathcal{L}^1 - \frac{1}{\delta} \int_{y-\delta}^{y} (f-g) d\mathcal{L}^1 = 0 \ .$$

Letting $\delta \to 0$, we get $f(x) - g(x) = f(y) - g(y)$, so $g = f + c$ for some constant c. □

6.6 BV Functions and AC Functions

Recall that a function f on $[a, b]$ is of bounded variation if there is some M such that

$$\sum_{j=1}^{n} |f(x_{j+1}) - f(x_j)| \le M$$

for all partitions $P: \ a = x_0 < x_1 < \cdots < x_n < x_{n+1} = b$ of $[a,b]$. Every monotone, real-valued function on $[a,b]$ is of bounded variation. For a function of bounded variation, its total variation (function) T_f is a function on $[a,b]$ given by

$$T_f(x) = \sup\left\{\sum_{j=1}^{n} |f(x_{j+1}) - f(x_j)| : \ a = x_0 < x_1 < \cdots < x_{n+1} = x \text{ on } [a,x]\right\}.$$

The total variation function is increasing and of bounded variation. As a result,

$$f_2 = \frac{1}{2}(T_f + f) \qquad \text{and} \qquad f_1 = \frac{1}{2}(T_f - f)$$

are also increasing. The decomposition $f = f_2 - f_1$ is called the Jordan decomposition of f.

Theorem 6.17. *Every function f of bounded variation on $[a,b]$ is differentiable almost everywhere. Furthermore, there is a signed Radon measure μ such that*

$$\int \varphi' f d\mathcal{L}^1 = -\int \varphi d\mu \ , \quad \forall \varphi \in C_c(a,b) \cap C^1(a,b). \tag{6.3}$$

Conversely, let $f \in L^1[a,b]$ and μ a signed Radon measure μ satisfy (6.3). Then f has a representative in $BV[a,b]$.

Proof. Let $f = f_2 - f_1$ be the Jordan decomposition of f. We extend f_1 and f_2 so that they are constant on $(-\infty, a]$ and $[b, \infty)$ and let $f = f_2 - f_1$. By modifying these functions on a null set we may also assume that they are left continuous. Let μ_1 and μ_2 be the finite Radon measures associated to $f_1 - f_1(a)$ and $f_2 - f_2(a)$ respectively. From (6.2) we see that

$$\int \varphi' f d\mathcal{L}^1 = -\int \varphi d\mu \ , \quad \forall \varphi \in C_c(\mathbb{R}) \cap C^1(\mathbb{R}) \ , \quad \mu = \mu_2 - \mu_1 \ ,$$

holds.

Conversely, let $\mu = \mu_2 - \mu_1$ be the Jordan decomposition of the Radon signed measure μ. we extend μ_1, μ_2, μ to $\mathbb{R}$ by setting $\mu_1(E) = \mu_1(E \cap [a,b])$, etc. Let f_1 and f_2 be the distribution functions of μ_1 and μ_2 respectively. Then

$$\int \varphi'(f - (f_2 - f_1)) d\mathcal{L}^1 = 0 \ , \quad \forall \varphi \in C_c[a,b] \ ,$$

implies $f = f_2 - f_1 + c$ for some c almost everywhere, see the proof of Proposition 6.16 for details. □

This theorem characterizes BV-functions as those integrable functions whose weak derivatives are Radon measures. In the following we pursue the further question: What are those BV-functions whose associated Radon measures are absolutely continuous with respect to the Lebesgue measure? The answer turns out simple; they are precisely the absolutely continuous functions. Traditionally this class of functions is characterized as those functions for which the fundamental theorem of calculus holds.

A function f on $[a,b]$ is called absolutely continuous if for every $\varepsilon > 0$, there exists a δ such that

$$\sum_k |f(x_{k+1}) - f(x_k)| < \varepsilon, \quad \text{whenever } \sum_k |x_{k+1} - x_k| < \delta$$

where $I_k = [x_k, x_{k+1}]$ are intervals in $[a,b]$ with mutually disjoint interior. The number of these intervals could be countable. An absolutely continuous function must be of bounded variation, but the converse is not always true.

Set T_f be the total variation function of f. It can be shown that T_f is absolutely continuous if f is absolutely continuous. As a result, both $f_2 = \frac{1}{2}(T_f + f)$ and $f_1 = \frac{1}{2}(T_f - f)$ are increasing and absolutely continuous.

Let $f \in BV[a,b]$. Letting $f = f_2 - f_1$ be its Jordan decomposition, we extend f, f_1 and f_2 to $\mathbb{R}$ by setting them to be constants on $(-\infty, a)$ and (b, ∞) respectively. Let $\mu = \mu_2 - \mu_1$ where μ_1 and μ_2 take f_1 and f_2 as their respective distribution functions. We will assume this in the following proof.

Theorem 6.18. *Let $f \in BV[a,b]$. The following statements are equivalent:*

(a) f is absolutely continuous on $[a,b]$.

(b) The associated Radon measure μ is absolutely continuous with respect to the Lebesgue measure.

(c) f' exists a.e., belongs to $L^1[a,b]$ and

$$f(x) = f(a) + \int_a^x f' \, d\mathcal{L}^1, \quad \forall x \in [a,b],$$

holds.

Proof. We can assume that f is increasing. The general case follows from the Jordan decomposition.

(a) $\Rightarrow$ (b). After the extension f is absolutely continuous on $\mathbb{R}$. For $\varepsilon > 0$, there is some $\delta > 0$ such that

$$\sum_k |f(b_k) - f(a_k)| < \varepsilon,$$

whenever (a_k, b_k)'s are pairwise disjoint and

$$\sum_k (b_k - a_k) < \delta \ .$$

Now, let E be a null set in Lebesgue measure. For $\delta > 0$, there are intervals (c_k, d_k) whose union covers E and satisfies $\sum_k (c_k - d_k) < \delta$. We can find pariwise disjoint (a_k, b_k), $k \geq 1$, such that

$$\bigcup_k (a_k, b_k) = \bigcup_k (c_k, d_k).$$

We have

$$\sum_k (b_k - a_k) \leq \sum_k (c_k - d_k) < \delta.$$

Therefore,

$$\begin{aligned} \mu(E) &\leq \mu\left(\bigcup_k (c_k, d_k)\right) \\ &= \mu\left(\bigcup_k (a_k, b_k)\right) \\ &= \sum_k \mu(a_k, b_k) \\ &\leq \sum_k \left(f(b_k) - f(a_k)\right) \\ &< \varepsilon \ , \end{aligned}$$

whence E is also μ-null.

(b) $\Rightarrow$ (c). From (6.2),

$$\int \varphi' f \mathcal{L}^1 = -\int \varphi f' d\mathcal{L}^1, \quad \forall \varphi \in C_c(\mathbb{R}) \cap C^1(\mathbb{R}).$$

Let $x, y, x < y$, be two Lebesgue points of f in (a, b). As in the proof of Proposition 6.16 we deduce from this identity

$$f(y) = f(x) + \int_x^y f' d\mathcal{L}^1 \ .$$

It follows that f is continuous and one can take $x = a$ and y any point in $[a, b]$.

(c)$\Rightarrow$ (a) Let us show that for every integrable function g,

$$G(x) \equiv \int_a^x g \, d\mathcal{L}^1,$$

is absolutely continuous on $[a, b]$. Recall the absolute continuity of Lebsegue integral: Letting $g \in L^1[a, b]$, for every $\varepsilon > 0$, there exists some δ such that

$$\int_E g\, d\mathcal{L}^1 < \varepsilon, \qquad \text{whenever } E \text{ is measurable and } \mathcal{L}^1(E) < \delta.$$

Taking $E = \bigcup_j I_j$ where I_j are disjoint intervals, we have

$$\mathcal{L}^1(E) = \sum_j \mathcal{L}^1(I_j) = \sum_j |x_{j+1} - x_j| < \delta,$$

which implies

$$\int_E g\, d\mathcal{L}^1 = \sum_j |G(x_{j+1}) - G(x_j)| < \varepsilon,$$

so G is absolutely continuous on $[a, b]$.

□

Comments on Chapter 6. We follow [EG] closely in this chapter except on the definition of the Radon measure. If you google for Radon measure, you will find there are different definitons. Here we use one that is consistent with our terminology in the last semester. Aside from the clarification of Radon measure there are very few changes. I do not copy some of the proofs. Please look up 1.5–1.7 in [EG]. The proof of Besicovitch's covering theorem is left out. You may google for more on this theorem.

There is another version of Vitali's covering theorem where the collection of balls is a finite one. Then the number 5 can be replaced by 3, see exercise and [R1].

Lebesgue points and sets were introduced by Lebesgue in the study of Fourier series. A celebrated theorem of his asserts that the Cesàro sums of the Fourier series of a Lebesgue integrable function converge to the function at every Lebesgue point. Look up chapter 5 of [HS] for a proof.

Reflections on the definition of Lebesgue points. As pointed out already, Lebesgue points vary from representatives of the same L^1-function. In the following we let $[f]$ denote the equivalence class of functions to which f, a pointwisely defined function, belongs. For instance, consider f to be the constant function 1 and f_E the function which is equal to 1 except on the Lebesgue null set E, where it is assigned to some values not equal to 1. Let $\mathcal{L}(f)$ be the Lebesgue set of f. We have $\mathcal{L}(f) = \mathbb{R}^n$ and $\mathcal{L}(f_E) = \mathbb{R}^n \setminus E$. So the concept of Lebesgue points is not intrinsic. It is possible to obtain an intrinsic definition of a Lebesgue point. Indeed, a point x is called an **intrinsic Lebesgue point** of a locally integrable function $[f]$ in $L^1(\mu)$ if there exists some $c \in \mathbb{R}$ such that

$$\lim_{r \to 0} \frac{1}{\mu(\overline{B}_r(x))} \int_{\overline{B}_r(x)} |f(y) - c| d\mu(y) = 0,$$

where f is any representative of $[f]$. It is clear that the value c, whenever exists, is the same for all representatives. Corollary 6.12 tells us that all intrinsic Lebesgue points form a set of full measure. Let us call it the **intrinsic Lebesgue set** of $[f]$, denoted by $\mathcal{L}([f])$. With this at our disposal, given $[f] \in L^1(\mu)$, we define its **precise representation** f_p to be, for $x \in \mathcal{L}([f])$, set $f_p(x)$ equal to the value c as in the above limit; and, for $x \notin \mathcal{L}([f])$, set $f_p(x) = \infty$. (You may also set it to be 0 so that f is finite everywhere.) Clearly, $f_p \in [f]$ and $\mathcal{L}(f_p) = \mathcal{L}([f])$. The precise representation of a locally integrable function has the nice property that $f_p(x)$ is finite if and only if

$$\lim_{r\to 0} \frac{1}{\mu(\overline{B}_r(x))} \int_{\overline{B}_r(x)} |f(y) - f_p(x)| d\mu(y) = 0, \quad \forall f \in [f].$$

Furthermore, the precise representation has the maximal property, namely, for any representative f of $[f]$, $\mathcal{L}(f) \subset \mathcal{L}(f_p)$. An interesting question is, given a null set E, is there some $[f] \in L^1_{\text{loc}}(\mu)$ so that $\mathcal{L}([f]) = \mathbb{R}^n \setminus E$? I don't know the answer even when $\mu = \mathcal{L}^1$ and E is Borel.

One may find a detailed discussion on BV- and AC-functions in Royden's Real Analysis which is assumed to be covered in the undergraduate real analysis. Our treatment on these functions differs from those in Royden's and [R1] so that its extension to higher dimensions becomes apparent. We have restricted things to a bound interval $[a, b]$ in order to compare with the old results. In general one can define $BV(\mathbb{R})$ and $AC(\mathbb{R})$. The Sobolev space $H^1(\mathbb{R}^n)$ consists of all L^1-functions whose weak derivatives f_k defined via

$$\int f \frac{\partial \varphi}{\partial x_k} d\mathcal{L}^n = - \int f_k \varphi d\mathcal{L}^n,$$

are also in $L^1(\mathbb{R}^n)$. It reduces to $AC(\mathbb{R})$ when $n = 1$. On the other hand, $BV(\mathbb{R}^n)$ consists of all L^1-functions whose weak derivatives are Radon measures. It reduces to $BV(\mathbb{R})$ when $n = 1$. The definition is essentially the same as the relation in (6.3), see chapter 5 in [EG] for a detailed treatment.